Geophysical Journal International

Geophys. J. Int. (2013) **192,** 1332–1341 doi: 10.1093/gji/ggs103

Locating a scatterer in the active volcanic area of Southern Peru from ambient noise cross-correlation

Yiran Ma, Robert W. Clayton, Victor C. Tsai and Zhongwen Zhan

Seismological Laboratory, California Institute of Technology, CA 91125*, USA. E-mail: yrma@caltech.edu*

Accepted 2012 December 5. Received 2012 November 3; in original form 2012 July 17

SUMMARY

We report on a strong scatterer of seismic energy in the 5–10 s period range located in the volcanic arc of Southern Peru. It is superficially like an active noise source in that it produces a continuous signal that arrives earlier than the inter-station surface wave in the noise cross-correlations. However, it is clearly determined to be a scatterer based on the coda arrivals observed in the cross-correlations, and the fact that it scatters waves from earthquake sources. We model the scatterer as a cylinder approximately 5 km in diameter with a shear wave velocity 30 per cent lower than the background velocity. It is likely to exist at the depth of 5–10 km, and is located at 71.6°W/16.1°S with an error of 10 km, which is near the inactive volcano Nevado Chachani and the active volcano El Misti which recently erupted in 1985.

Key words: Interferometry; Wave scattering and diffraction; Volcanic arc processes.

1 INTRODUCTION

Ambient noise cross-correlations have been used extensively in the past decade to estimate the Green's function between pairs of stations (Lobkis & Weaver 2001; Shapiro & Campillo 2004; Wapenaar 2004). The surface wave portion of the Green's function has been used to determine earth structure (Shapiro *et al.* 2005; Yao *et al.* 2006; Brenguier *et al.* 2007; Lin *et al.* 2008) and, by looking at time-lapse changes, the noise cross-correlations have been used to forecast volcanic eruptions (Brenguier *et al.* 2008). The primary noise sources are primary microseisms (Hasselmann 1963) in the 10–20 s period band and secondary microseisms (Longuet-Higgins 1950) in 5–10 s period band, but unidentified sources such as the 26-s microseismic energy originating near Africa (Shapiro *et al.* 2006) have also been reported. When noise sources are non-uniformly distributed, Green's functions can still be approximately recovered (Yang & Ritzwoller 2008), but artificial or truncation phases can be created (e.g. Zhan *et al.* 2010).

Recently, Zeng & Ni (2010) and Zheng *et al.* (2011) reported on precursors which arrive earlier than the inter-station surface wave in ambient noise cross-correlations. They attribute the precursors to a localized noise source and locate it on Kyushu Island, Japan, by a similar procedure to the one described in this paper. Zeng & Ni (2011) suggest that those signals are created by random events associated with Aso volcano. In this case, Aso volcano is thought to be an active noise source because the strength of the precursors is out of phase with the seasonal changes of oceanic noise, and local measurements (Kawakatsu *et al.* 1994) show that Aso volcano produces significant energy in the frequency range of the precursors.

In this paper, we report on a situation that is superficially like that of the Zeng & Ni (2010, 2011) study. Strong precursory arrivals are seen on noise cross-correlations in southern Peru, and the source of this energy is located in the volcanic arc associated with regional subduction (see Fig. 1). However, as we will show, the source in this case is not an active source, but a passive scatterer of oceanic microseism energy. This conclusion is confirmed by observing that the inferred scatterer also scatters waves from earthquakes. This type of feature may cause unidentified arrivals that appear on noise cross-correlations in other studies.

2 PRECURSOR AND CODA WAVES

Using the method of Bensen *et al.* (2007), we perform vertical–vertical ambient noise cross-correlations between all pairs of broadband stations along the lines shown in Fig. 1. A profile of cross-correlations from station PE13 to all stations on line PE (blue dots, approximately perpendicular to the coast) is shown in Fig. 2(a), and a profile of cross-correlations from station PF25 to all stations on line PF (green dots, approximately 300 km inland and parallel to the coast) is shown in Fig. 2(b). Each trace is filtered to 5–10 s period band and normalized by its maximum amplitude and, on average, 2 yr of recordings are used in each cross-correlation. The direct surface wave arrivals (red lines in Fig. 2) propagating inland are clear in Fig. 2(a), but it is apparent from both of these profiles that the cross-correlations contain significant energy in arrivals besides the direct surface wave. The precursory energy (arriving at correlation times closer to zero than the direct wave) in both profiles is especially clear, but we observe that significant energy also arrives later than the direct wave (e.g. at stations PE03-PE12). We devote the remainder of this paper to quantitatively explaining these observations.

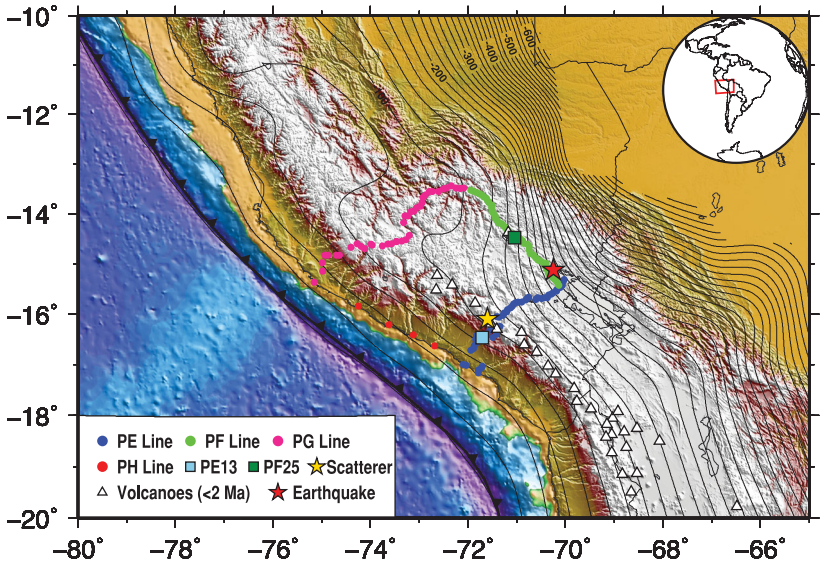

Figure 1. Locations of the seismic stations (dots) used in this study: lines PE (blue), PF (green), PG (pink) and PH (red). Stations PE13 (light blue square) and PF25 (dark green square) are the two virtual source stations used in Section 2. The yellow star is the location of the scatterer located in Section 3. The red star is the hypocenter of the earthquake used in Section 5. All the white triangles are volcanoes (<2 Ma).

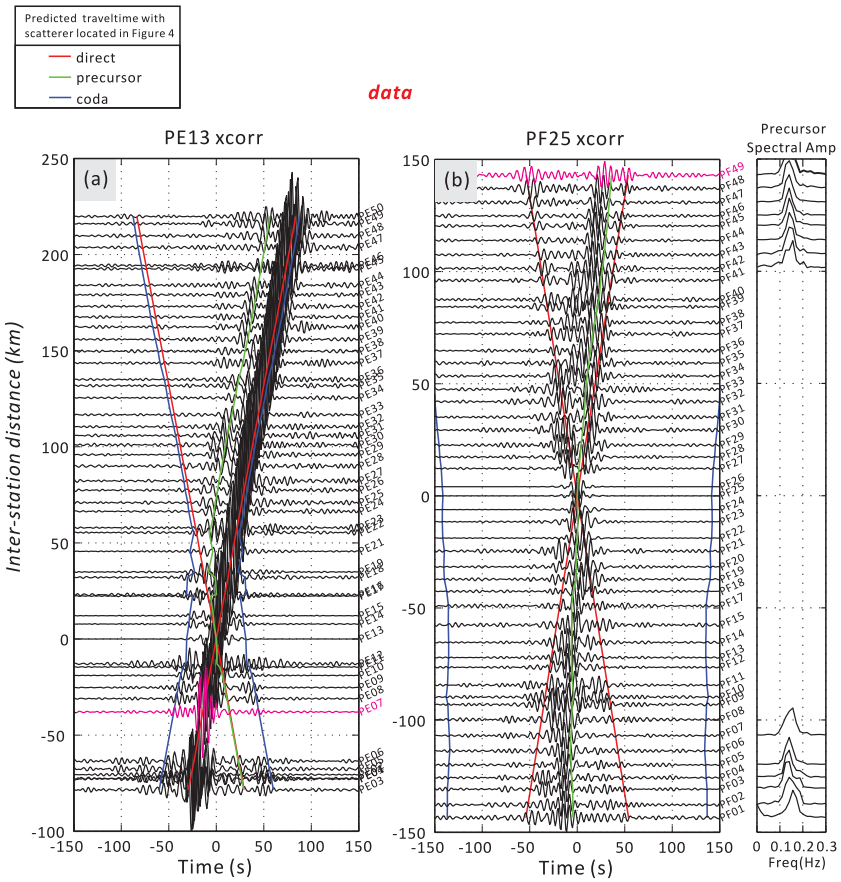

Figure 2. Two profiles of cross-correlations. The direct, coda and precursor arrival times plotted in red, blue and green lines are predicted from the location of the scatterer shown in Fig. 1 and located in Section 3. (a) Cross-correlations between PE13 (the light blue square in Fig. 1) and all PE stations (the blue dots in Fig. 1), aligned in a direction roughly perpendicular to the coast. Positive distances represent the cross-correlations between PE13 and stations further inland. The pink trace is the cross-correlation between PE13 and PE07, which is examined in Section 4. (b) Cross-correlations between PF25 (the dark green square in Fig. 1) and all PF stations (the green dots in Fig. 1), aligned in a direction roughly parallel to the coast. Positive distances represent the cross-correlations between PF25 and stations to the NW. The pink trace is the cross-correlation between PF25 and PF49, which is examined in Section 4. Also shown are some examples of the spectral amplitude of the precursor. The spectral amplitudes of the precursory waves are peaked at about 0.15 Hz.

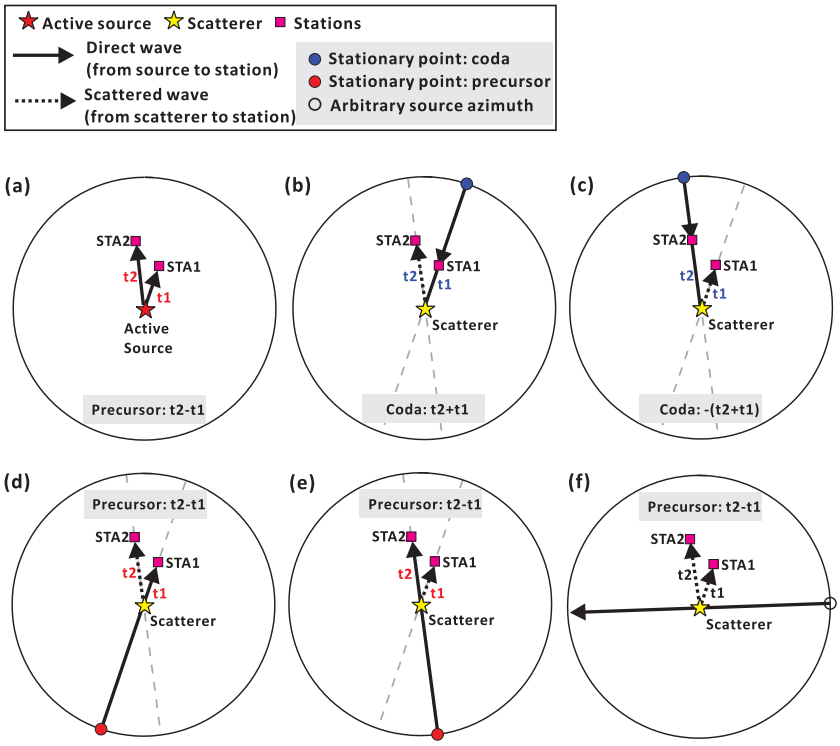

Figure 3. Schematic diagram shows the generation of the precursor and coda in the cross-correlation. (a) In the case of an active source, the precursor is generated from the cross-correlation between two waves travelling from the source to the two stations. (b)–(f) In the case of a scatterer, both coda and precursor are generated. In (b)–(e), the precursor and coda are generated from the 'direct-scattered' wave correlations. Among the noise sources from all azimuths, the contributions are mainly from four stationary points denoted with coloured dots. In (f), the precursor is generated from the cross-correlation between two scattered waves, and no stationary point exists.

We first focus on understanding the precursors. There are two clear candidates for explaining the precursory energy. Either the precursors are due to the presence of an active source of microseismic waves or they are due to the presence of a passive scatterer. In either case, if the travel times of the surface wave from the source/scatterer location to a pair of stations STA1 and STA2 are t_1 and t_2, then the cross-correlation between STA1 and STA2 produces an additional signal that arrives at $t_2 - t_1$. This arrival will always be precursory to the direct arrival between the stations except when the stations and source/scatterer are co-linear. For the active source, this is straightforward to understand and a simple schematic is shown in Fig. 3(a).

For the scatterer, we can understand all of the arrivals based on the stationary phase analysis of Snieder *et al.* (2008), who derived the travel times and stationary points for the expected arrivals in the cross-correlation of ambient noise in the presence of a point scatterer. Wapenaar *et al.* (2010) also derived similar relations using the reciprocity theorem. Their results show that the cross-correlation contains 'direct-scattered' wave correlations (Figs 3b–e). These 'direct-scattered' correlations represent the cross-correlation between the direct wave that propagates from the source to one station and the scattered wave from the source to the scatterer then to the other station. These contributions, which come from four stationary points, result in four arrivals. If t_1 and t_2 are the travel times of the surface wave from the scatterer to the two stations STA1 and STA2, then two of the four arrivals emerge as coda waves in the positive and negative lags with the arrival times $\pm(t_1 + t_2)$ (Figs 3b and c) and the other two are precursor waves and have the same arrival time $t_2 - t_1$ (Figs 3d and e). The precursor also coincides with the arrival time of the signal from the correlation between two scattered waves that propagate to the two stations (Fig. 3f). If sources are uniformly distributed, the three components of the precursor wave cancel each other (Snieder *et al.* 2008), but when the noise source distribution is not uniform or the wave speeds vary spatially, the phase exists.

From the preceding discussion, we find that while both the active source and the scatterer generate precursors, coda waves are only generated by the scatterer. Since strong coda waves are observed in Fig. 2(a), we rule out the possibility that an active source is exclusively responsible for the observations. In the following sections, we determine the location (Section 3) and strength (Section 4) of the inferred scatterer, as well as provide a more detailed analysis of the features in Fig. 2 and offer additional evidence against an active source. In Section 5, we confirm that the inferred scatterer also scatters earthquake waves and therefore cannot be a microseismic source.

3 LOCATING THE SCATTERER

We use all the cross-correlations between PE stations and PF stations to locate the scatterer. If the scatterer is equidistant from a given pair of stations, the precursor will arrive at zero lag assuming uniform velocity. Searching for pairs of stations which have the precursor at zero lag allows us to quickly locate the scatterer to the active volcanic area. We then refine the location with a grid search using a $0.05^\circ \times 0.05^\circ$ grid. For each point, we predict the precursor and coda arrival times and stack the envelope amplitude of the cross-correlations (filtered to 5–10 s period band) at these times. The result is shown in Fig. 4. We use an average group velocity of

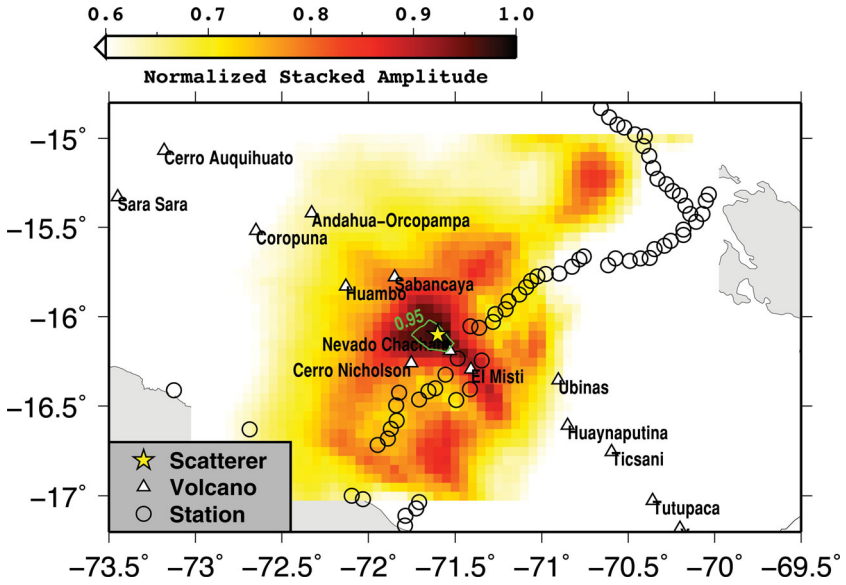

Figure 4. Locating the scatterer using a $0.05° \times 0.05°$ grid. For each point, we calculate the precursor and coda arrival times and stack the envelope amplitudes of the cross-correlations at these times. The colour is the normalized stacked amplitude for the entire data set, with arrival times calculated with a group velocity of 2.7 km s^{-1}. The 95 per cent confidence interval is denoted with a green contour, and is estimated through the bootstrap method. The scatterer's best-fit location is marked with a yellow star (71.6°W/16.1°S). It is closest to the inactive volcano Nevado Chachani, and is about 30 km from the active volcano El Misti which recently erupted in 1985. The locations of the volcanoes are from Global Volcanism Program (http://www.volcano.si.edu/world/).

2.7 km s^{-1}, obtained from 6-s direct surface wave group velocity measurements (as in Bensen *et al.* (2007)) between all stations of the array. The 6-s group velocity of the backarc region, where most of the cross-correlations used in the locating are, ranges from 2.6 to 2.9 km s^{-1}. A constant velocity for location purposes is justified because the velocity range is small and the paths average the velocity. Amplitudes are not used if they lie in the direct wave window defined as 10 s before the arrival time calculated with 3.0 km s^{-1} and 10 s after that calculated with 2.5 km s^{-1}. The maximum stacked amplitude and hence the best-fit location of the scatterer is at 71.6°W/16.1°S (yellow star in Figs 1 and 4). We note that we did not use the cross-correlations from lines PH and PG because we did not observe clear precursor and coda arrivals in the cross-correlations for these two lines. The reason for this is probably that line PG is far from the scatterer and line PH only has four stations, which makes it hard to observe the phases that are continuously observed in lines PE and PF (which each have 50 stations).

We estimate the 95 per cent confidence interval of the location using the bootstrap method. We vary the average velocity from 2.5 to 2.9 km s^{-1} in 0.1 km s^{-1} steps. For each velocity, we use 2000 data randomly picked from the entire 4600 cross-correlation data set to calculate the best-fit location, and repeat this 20 000 times. All the best-fit locations are combined to produce a contour map of the probability. In Fig. 4, the 95 per cent confidence interval contour is shown in green.

With the best-fit location of the scatterer, we then predict the travel times of the precursor and coda waves using the 2.7 km s^{-1} group velocity. These are plotted as coloured lines in Fig. 2, and they fit the group arrivals, which are the peak of the envelopes, to within 1.5 wavelengths. For example, the continuously observed precursor wave seen in Fig. 2(a) from PE19-PE50 is well fit by the green synthetic line. The coda waves observed from PE03-PE19 are also well fit by the blue synthetic lines. We note that the predicted arrival times do not just increase linearly with distance (as they would if plotted as a function of scatterer-station distance) because the distances plotted in Fig. 2 are inter-station distances. Some predicted arrivals do not show up clearly in the data because of the non-uniform source distribution. For example, the cross-correlations in Fig. 2(a) are always one sided because of the dominant noise sources from the Pacific. Since the cross-correlations in Fig. 2(b) are roughly parallel to the coast, the direct waves are not well illuminated by oceanic noise sources, making the precursor wave the strongest arrivals in these cross-correlations. In the next section, these amplitude differences are modelled.

4 MODELLING THE PRECURSOR AND CODA WAVE AMPLITUDES

As discussed previously, because of the observed coda, scattering must contribute to the generation of the precursor. While an active source may also exist, the following two lines of evidence suggest that it should not be an important factor in producing the precursors. Firstly, the frequency content of the precursor is peaked in the secondary microseism band (peak at 0.15 Hz), as is shown in Fig. 2 for station pairs on line PF with especially clear precursory arrivals. This agreement suggests that the energy is related to oceanic sources. A scatterer of oceanic noise would naturally have this feature, whereas the coincidence of peak amplitude for an active source would need additional explanation. Secondly, as with the PE13 profile, we observe a sudden decrease in the precursor amplitude when STA2 changes from inland of PE19 (PE19-PE50) to the coastal side of PE19 (PE03-PE19). This cannot be explained by an active source unless the source has a strong directivity, but it is well predicted by the stationary points of the precursors that will be discussed in Section 4.3.

Based on the evidence presented above, we believe that the scatterer alone can explain our observations, and we proceed to model the precursor and coda amplitudes with a single scatterer. The amplitudes of the precursor and coda waves depend on the azimuthal distribution of noise source strength and the scatterer properties (e.g. velocity, size and depth). To fit the cross-correlation amplitudes, the first step is to determine the azimuthal variations in the

ambient noise source strength by using the direct wave amplitudes in the cross-correlations at different azimuths.

4.1 Noise source strength

The noise source strength at azimuth θ, written as $S(\theta)$, can be determined by the direct wave amplitude since $S(\theta)^2 \propto A\sqrt{R}$, where A is the direct wave amplitude and R is the inter-station distance (Stehly *et al.* 2006; Yang & Ritzwoller 2008; Yao *et al.* 2009). Here, we do not consider attenuation effects since the path lengths inside the array are short. We use 1-yr cross-correlations of all the stations with at least 300 days' recording in the year 2011. The cross-correlations are filtered to 5–10 s period band. We then calculate the signal-to-noise ratio (SNR) which is defined as the ratio of the peak of the envelope in the signal window (corresponding to a velocity between 2 and 3.5 km s^{-1}) to the root mean square (RMS) of the noise window from 1000 s after the signal window to the end of the cross-correlation (3000 s lag). The SNR is multiplied by the square root of the inter-station distance to account for the geometrical spreading of the Rayleigh wave, and the (normalized) source strength is the square root of the result. To smooth the measurements, we average the source strength over every 10° in azimuth. The above method of using SNR to quantify the source strength works because the same normalization procedures are applied to both the signal and the noise window in the cross-correlation, and the trailing noise level is assumed to be similar for the array (Yang & Ritzwoller 2008; Lin *et al.* 2011).

We expect that the peak of the envelope samples the group arrival of the direct wave, and one problem is to avoid sampling the precursor if its amplitude exceeds that of the direct wave. To address this problem, we first visually inspect all the cross-correlations plotted with the predicted travel times of the precursor and direct waves.

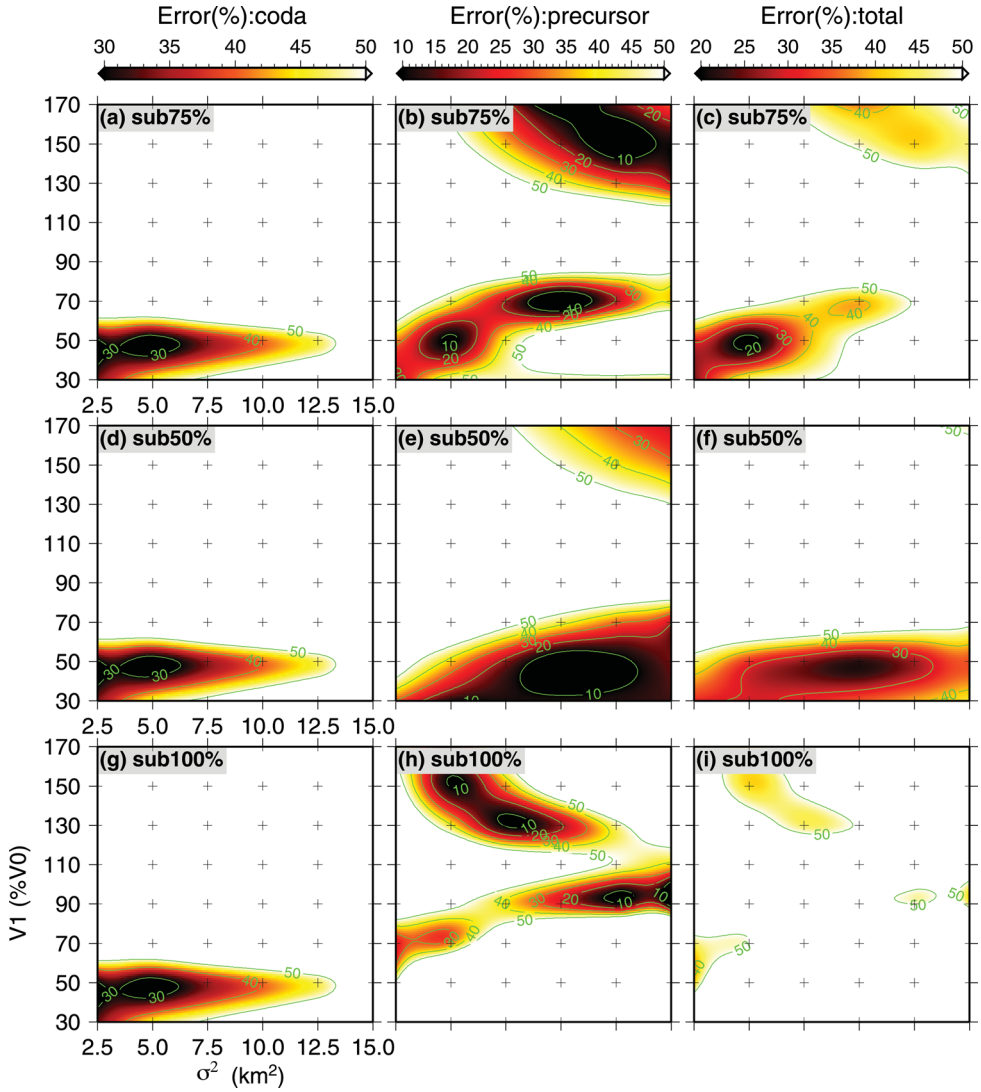

Figure 5. Modelling results for all combinations of parameters σ^2 and v_1 (see eq. 2 for definitions). (a) The error in fitting the coda amplitudes in the cross-correlation between PE13 and PE07 (pink trace in Fig. 2a). (b) The error fitting the precursor amplitude in the cross-correlation between PF25 and PF49 (pink trace in Fig. 2b). (c) The average of (a) and (b). The error is the relative error defined as $|A_{syn} - A_{obs}|/A_{obs}$. The three rows (sub 75 per cent, sub 50 per cent and sub 100 per cent) show the errors under three modified source conditions (see the text for details), which all give the same optimal parameters of σ^2 =5.0 km^2 and v_1 = 50 per cent v_0 from coda fitting. With the optimal scatterer properties, the precursor is best fit under the first (sub 75 per cent) source condition.

For those cross-correlations with a comparatively strong precursor, we calculate the precursor window using the same velocity range as that for the direct wave window. We also empirically determine the maximum SNR for the precursor as 40. If the calculated SNR is above this value, it is assumed to sample the direct wave even if the precursor window and the direct wave window overlap. If the calculated SNR is less than 40 and the windows overlap, the peak of the envelope is likely to sample the precursor and we determine the timing of the direct wave from the reciprocal lag since it is not contaminated by the precursor. We accept the timing of the direct wave determined in the reciprocal lag if the corresponding source strength is larger than 10, and use this timing to sample the envelope amplitude in the original lag and calculate the SNR.

We note that in the azimuthal source strengths estimated from the SNRs, the weak sources are heavily overestimated because they contain a significant component of background waves generated by scattering from throughout the volume, and the correlation of uncorrelated noise at each station site. To model the amplitude of the "weak" direct waves, we correct this bias by removing an estimate of the background wave strength to obtain a new estimate of source strength $S'(\theta)$. As will be discussed in Section 4.3, it is important to note that this correction only affects the modelled amplitudes of the direct waves and does not affect the estimation of scatterer properties.

4.2 Modelling the scatterer

To test the size and velocity contrast of the scatterer, we use a 2-D finite-difference code (Li *et al.* in preparation) to do membrane wave modelling. Tanimoto (1990) showed that for a narrow frequency band, and for smooth lateral variations of elastic constants (Lame parameter λ and shear modulus μ), Rayleigh and Love waves can be approximated by membranes waves which satisfy the 2-D wave equation:

$$\frac{\partial^2 u}{\partial t^2} - c\,(x, y)^2\,\nabla^2 u = 0, \tag{1}$$

where u is the displacement and c is the local Rayleigh or Love wave phase velocity.

The model we use is a uniform phase velocity model of $v_0 =$ 3 km s^{-1} with an embedded scatterer. The phase velocity is the average of that measured from the direct surface wave (method by Yao *et al.* (2006)) at 6 s in the cross-correlations between stations of this array. The scatterer is assumed to have a Gaussian distribution of velocity described by:

$$v = v_0 + (v_1 - v_0) \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right), \tag{2}$$

where r is the distance from the centre of the scatterer, v_1 is the velocity at the centre and σ is one standard deviation of the Gaussian

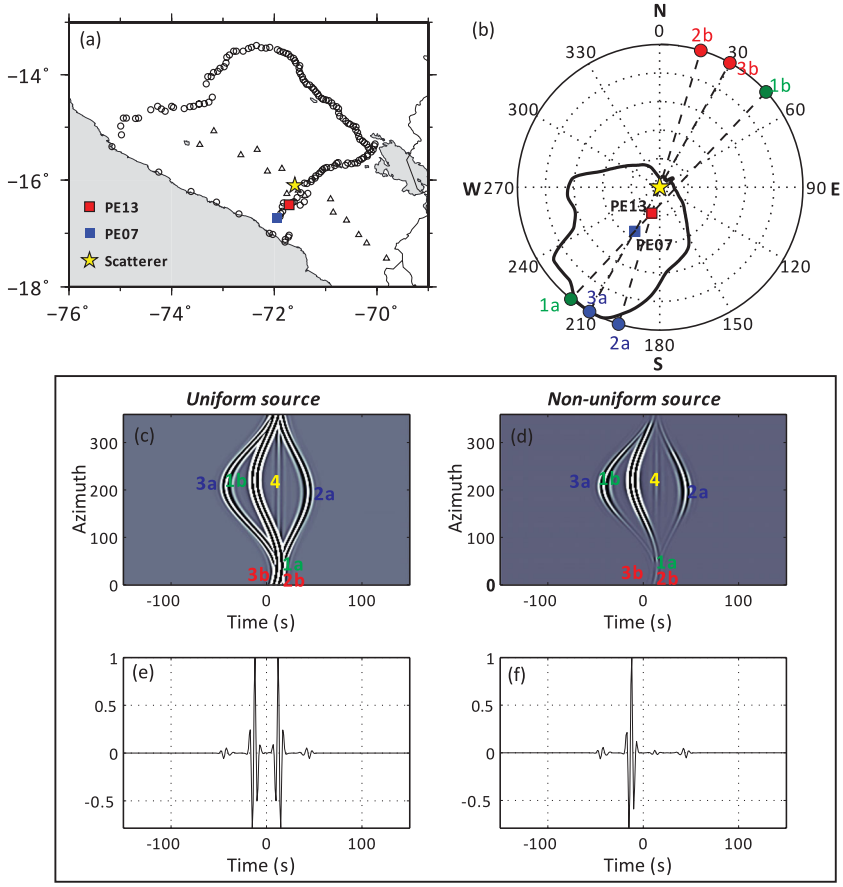

Figure 6. Modelling the cross-correlation between PE13 and PE07. The parameters used are: $\sigma^2 = 5.0$ km^2 and $v_1 = 50$ per cent v_0, where $v_0 = 3$ km s^{-1} is the background velocity. (a) Locations of the two stations and the scatterer. (b) A polar diagram showing the location of the stationary points of the cross-correlation. The black line is the source strength at different azimuths, and is strongest toward the coast and weakest in the opposite direction. (c) The cross-correlogram showing the cross-correlation for each source individually. The y-axis is the azimuth of the source. The colour represents the amplitude of the cross-correlation. The traces with stationary points are described in the text. (d) Same as (c), but for the non-uniform source case using the realistic source strength in (b). (e) The final cross-correlation from stacking all of the individual cross-correlations in (c). (f) The final cross-correlation from stacking all of the individual cross-correlations in (d). The synthetic result shown in (f) is similar to the data shown in pink in Fig. 2(a).

function. We take the cross-correlations between PE13 and PE07 and between PF25 and PF49 as two samples, and fit the precursor and coda amplitudes. We perform synthetics for a range of values including $\sigma^2 = 2.5$, 5.0, 7.5, 10.0, 12.5 and 15.0 km^2 and v_1 =30, 50, 70, 90, 110, 130, 150 and 170 per cent v_0.

A circular distribution of noise sources centred at the scatterer is assumed. The sources are placed 800 km from the scatterer so that they are in the far field compared with the scatterer–station distance. The source wavelet is a Ricker wavelet with a central period of 6 s. After calculating the waveforms recorded at the two stations from each source, we perform cross-correlations between the recordings for each source individually and stack them to get the final cross-correlation result. For the non-uniform source case, we weight the individual cross-correlations by the square of the source strength ($S'(\theta)$).

4.3 Modelling results

Fig. 5 shows the misfits for the coda amplitudes in the cross-correlation between PE13 and PE07 (pink trace in Fig. 2(a)), and the precursor amplitude in the cross-correlation between PF25 and PF49 (pink trace in Fig. 2(b)). As mentioned previously, our SNR estimate of source strength ($S(\theta)$) is an overestimate due to significant background scattering and uncorrelated noise, and we correct for this by subtracting a given percentage of the minimum $S(\theta)$ to create our revised estimate$S'(\theta)$. Specifically, we set $S'(\theta) = S(\theta) - \xi \min[S(\theta)]$, where ξ is either 75, 50 or 100 per cent. Synthetics for these three modified source conditions are plotted in the three rows of Fig. 5.

We see that the coda misfits are mostly unaffected by the choice of source modification, with all three conditions giving the same optimal parameters of σ^2 =5.0 km^2 (i.e. σ = 2.2 km) and v_1 = 50 per cent v_0 with about 27 per cent error. We note that this optimal Gaussian scatterer has a velocity perturbation of 30 per cent v_0 at one-sigma (2.2 km) radius. The coda fits are largely unaffected because, as shown in Fig. 6(b), coda amplitudes are primarily determined by sources close to the maximum in $S'(\theta)$ (towards the coast), which is nearly identical to $S(\theta)$ regardless of choice of ξ. The precursor misfits, though, are strongly affected by the modification. The precursor amplitude is related to the difference between the source strengths at the two stationary points since their contributions counteract each other. This difference is not affected if a constant value is subtracted from each. However, here, we try to fit the ratio between the precursor and direct wave amplitude. As shown in Fig. 7(b), for PF25-PF49, the stationary point of the direct wave is no longer in the very strong coastal direction, and thus is easily affected by the subtraction. With the optimal scatterer properties determined above, we find that ξ = 75 per cent fits the precursor best. Next, we will show the detailed modelling of the coda and precursor with our best fitting model, σ^2 =5.0 km^2, v_1 = 50 per cent v_0 and ξ = 75 per cent.

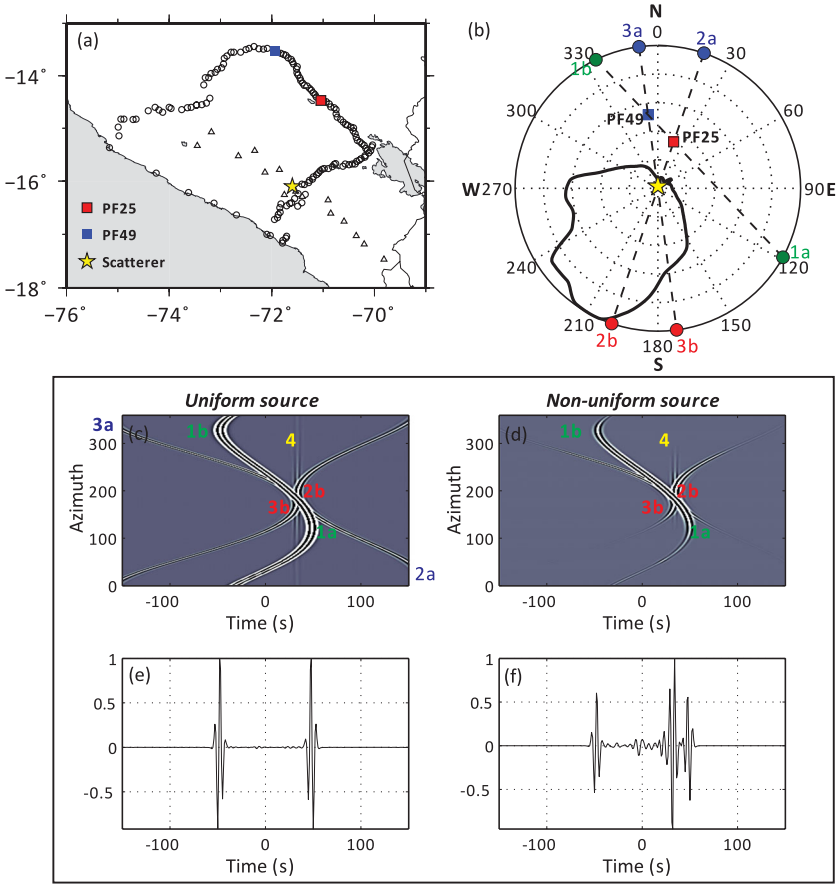

Figure 7. Modelling the cross-correlation between PF25 and PF49. Panels are as in Fig. 6, and use the same model parameters. We see that the precursor wave does not show up in the stacked cross-correlations, which means the contributions from 2b, 3b and 4 in (c) cancel each other. (f) The final cross-correlation from stacking all of the individual cross-correlations in (d). The precursor shows up in this non-uniform source case. This synthetic result in (f) is similar to the data shown in pink in Fig. 2(b).

Fig. 6 shows the simulated cross-correlation between PE13 and PE07. The location of the two stations and the scatterer is shown in Fig. 6(a). The source strength (thick black line) and stationary points (dots) are shown in the polar diagram (Fig. 6(b)). Both the uniform source and non-uniform source cases are shown. Figs 6(c) and (d) show the cross-correlations for each source individually, and Figs 6(e) and (f) are the final cross-correlations that result from stacking the individual cross-correlations in Figs 6(c) and (d), respectively. For both the uniform source and non-uniform source cases, the final cross-correlations are mainly from the four contributions labelled in Figs 6(c) and (d), respectively. Contribution 1 is from waves that directly travel from source to PE13 and PE07, and has two stationary points 1a and 1b corresponding to the two direct waves. Contribution 2 is the cross-correlation between one wave directly from the source to PE13 and one wave from the source to the scatterer then to PE07, and it has two stationary points for which 2a corresponds to the coda in the positive lag and 2b corresponds to the precursory arrival. Contribution 3 is the cross-correlation between one wave from the source to the scatterer and then to PE13 and one wave directly from the source to PE07, and it also has two stationary points for which 3a corresponds to the coda in the negative lag and 3b corresponds to the precursory arrival. Contribution 4 is the cross-correlation between two scattered waves and corresponds to the precursory arrival. We observe that the non-uniform case (Fig. 6f) fits the data well (*cf.*, the pink trace in Fig. 2a).

While the precursor is difficult to distinguish from the direct wave in the positive lag in the first example, they are well separated in the simulated cross-correlation between PF25 and PF49, as shown in Fig. 7. In Figs 7(c) and (d), which are the cross-correlations for each source individually, we see clearly that the stationary points 2b and 3b as well as contribution 4 all correspond to the precursor arrival time. In the uniform source case, the final stacked cross-correlation (Fig. 7e) shows no precursor signal, because the contributions from 2b, 3b and 4 cancel out, as predicted by theory (Snieder *et al.* 2008). However, for the non-uniform source case, the precursor does clearly emerge. The polar diagram (Fig. 7b) shows that the source strength corresponding to 2b and 3b differs significantly, which means the contributions cannot cancel.

We also produce synthetics for the two profiles shown in Fig. 2. The results are shown in Fig. 8, and are plotted with the same scale as the data in Fig. 2. The travel times are predicted with the phase velocity of 3 km s^{-1} used in the synthetics. Both sets of synthetics produce envelope amplitudes that fit the data reasonably well (*cf.*, Fig. 2). The variation in the amplitude of the precursor and coda is the same as that in the data. In general, the coda amplitude is related to the source strength at its stationary point (Fig. 3b for positive lag and Fig. 3c for negative lag). The precursor amplitude is related to the difference between the source strengths at its two stationary points (Figs 3d and e) since their contributions counteract each other as stated in Section 2 and modelled in Section 4 (Fig. 7).

For the PE13 profile (Fig. 8a), the scatterer is near co-linear with the PE line (stations numbered from coastal to inland) and is approximately at PE19 (see Fig. 1). The stationary point of the coda in the positive lag is in the direction pointing from the scatterer to PE13, and is in the coastal direction where the noise source is the strongest. Consequently, the coda in the positive lag should always be observable in PE13 profile, except for the distance range of PE19-PE50, where the positive coda travel time is so close to the direct wave that it is hidden by the direct wave. The stationary point of the coda in the negative lag is in the direction pointing from the scatterer to STA2, which is in the strong-source coastal direction when STA2 is in the range of PE03-PE19, and is in the weak-source inland

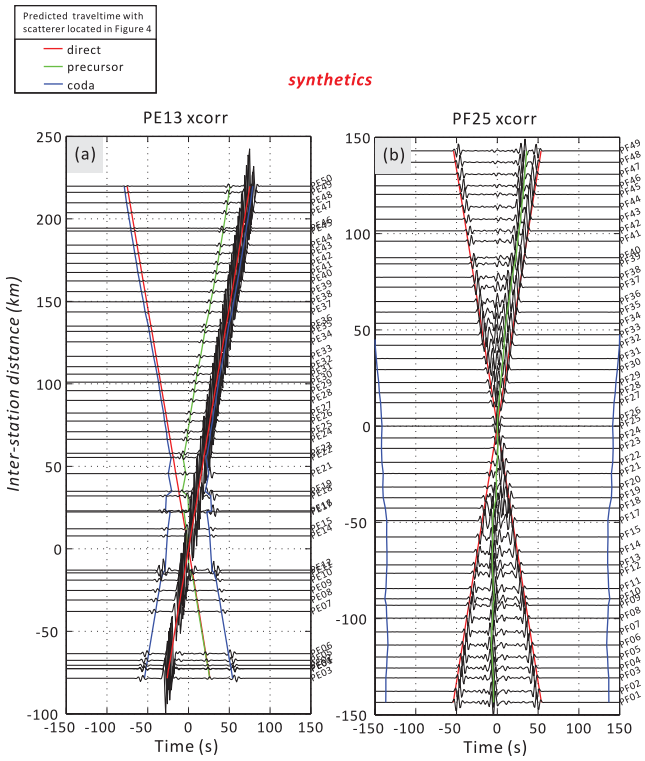

Figure 8. Modelling the cross-correlations between (a) PE13 and all PE stations, and (b) PF25 and all PF stations. The parameters used are: $\sigma^2 = 5.0$ km^2 and $v_1 = 50$ per cent v_0, where $v_0 = 3$ km s^{-1} is the background velocity. The synthetics match the envelope amplitude of the data in Fig. 2 reasonably well. See the text for the discussion of the variations in the precursor and coda amplitudes.

direction when STA2 is in the range of PE19-PE50. Therefore, the coda in the negative lag becomes invisible when STA2 changes from coastal to inland of PE19. For the precursor, when STA2 is in the range of PE03-PE19, the two stationary points are both in the inland direction and are so close that the difference in their source strengths is very small (e.g. Fig. 6b). As a result, the precursor is weak. However, when STA2 is in the range of PE19-PE50, the two stationary points are in the opposite directions (coastal and inland), which results in a strong precursor, as observed in Fig. 2(a).

For the PF25 profile (Fig. 8b), the coda stationary points for each station pair in this profile are always in the inland directions where the noise sources are weak, and therefore are always unrecognizable. However, the precursor stationary points are well separated and the source strengths differ significantly (e.g. Fig. 7b), and therefore the precursors are always strong in this profile, as observed in Fig. 2(b).

Since we use membrane wave modelling, we cannot directly determine the vertical size of the scatterer. Surface wave (group and phase) velocities are determined by an integral over structure with depth, but are most sensitive to shear velocity structure at depths of about one-third of their wavelength (Weeraratne *et al.* 2003; Yang *et al.* 2007). Since we use cross-correlations between 5 and 10 s period and phase velocities are about 3 km s^{-1} at 6 s in this region, we can deduce that the scatterer likely exists at depths of 5–10 km, where the velocity perturbation required is the smallest. If the scatterer were only at much shallower or deeper depths (e.g. $\mp$5 km), it would require a much larger velocity contrast, which we consider to be unlikely. However, we do not exclude the possibility that the scatterer extends much deeper than 10 km. We also note that we have not tested non-Gaussian scatterer models, which may

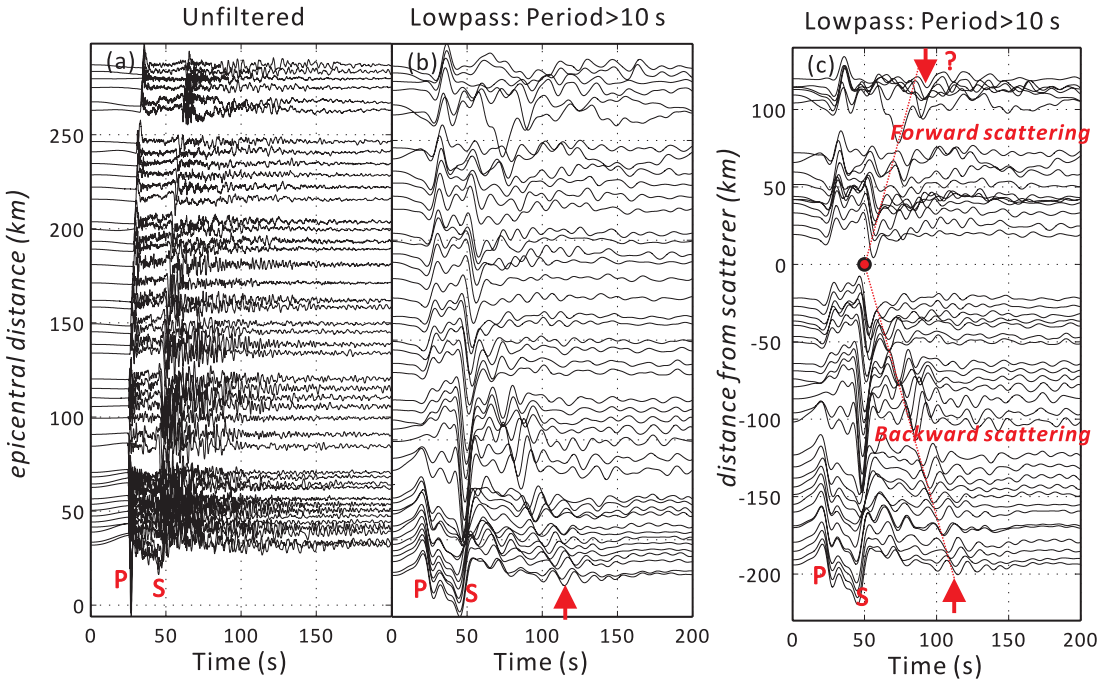

Figure 9. Vertical-component seismograms showing a clear backscattered wave. The hypocenter of the earthquake is the red star in Fig. 1. The epicentral depth is 210.2 km and the magnitude is 5.8. (a) Unfiltered data show clear *P* and *S* waves. (b) Low-pass filtered to period larger than 10 s. Clear backscattered waves appear, marked with the red arrow. (c) Traces sorted by the distance to the scatterer (yellow star in Fig. 1). The positive direction represents stations in the forward scattering region and the negative direction represents those in the backward scattering region. The backscattered wave shows a linear increase of travel time with distance and a velocity of about 3 km s^{-1}, indicating that it is a surface wave. The arrival time of the scattered wave at 0 km distance coincides with the *S* wave, which indicates that the energy is from scattering of the *S* wave. The forward scattered wave can also be observed though not as clearly.

also fit the data reasonably well but would have different model parameters.

5 CONFIRMATION BY AN EARTHQUAKE SOURCE

The scattering results shown above are confirmed by an earthquake recording with a clear backscattered wave (Fig. 9). This earthquake had an epicentral depth of 210 km and a magnitude of 5.8, and its hypocenter is shown as a red star in Fig. 1. The unfiltered (vertical component) seismograms (Fig. 9a) show clear *P* and *S* waves, and when they are low-pass filtered to periods longer than 10 s (Fig. 9b), we see a clear backscattered wave. In Fig. 9(c), we sort the traces according to distance from the scatterer, and distinguish the backward and forward scattering by negative and positive distances. We observe that the travel time of the backscattered wave is now a linear function of distance, and the velocity is about 3 km s^{-1}, which indicates that it is a surface wave. At zero offset, the scattered wave is coincident with the *S*-wave arrival, which indicates that the energy is from scattering of the *S* wave. The generation of a scattered surface wave indicates that the scatterer is shallow, consistent with our preferred scattering model of Section 4.

6 DISCUSSION AND CONCLUSIONS

We have determined that the strong precursor and coda arrivals that appear in the ambient noise cross-correlations within an array in Southern Peru are due to a scatterer located at 71.6°W/16.1°S with an error of about 10 km. The scatterer can be satisfactorily modelled as a vertical cylinder with a Gaussian distribution of velocity, but it is only required at depths of 5–10 km. The velocity at the centre of the Gaussian scatterer is 50 per cent of the background velocity, and at one-sigma (2.2 km) radius is 30 per cent lower than the background velocity. The scattering effects are confirmed by examining a local earthquake.

The scatterer is located within the volcanic arc of the subduction zone in Southern Peru and, as such, is likely related to volcanic processes. For example, it could be a low-velocity magma chamber. Nagaoka *et al.* (2012) imaged the magma chamber beneath an active volcano in Japan using ambient noise cross-correlations. Their phase velocity map (Fig. 3d therein) at periods of 5–10 s shows a ∼20 per cent lower velocity anomaly with radius of about 5 km. The *S*-wave velocity anomaly is about 5–10 km deep, and is ∼8 km offset from the nearby volcanic edifice. The features of their magma chamber are quite similar to the scatterer discussed here. Our best-fit location of the scatterer is closest to (∼10 km away from) the volcano Nevado Chachani which is presently inactive, and is about 30 km from the active volcano El Misti. Considering our location errors, the scatterer could be related to either of these volcanoes. The scatterer is close to but approximately 50 km outside of the Hualca Hualca inflation anomaly reported by previous InSAR studies (Pritchard & Simons 2002, 2004). Note that the inflation stopped in 1997 during their study, which is perhaps related to a large eruption of nearby Sabancaya volcano (Pritchard & Simons 2004).

This study points out that scatterers can generate coherent arrivals on ambient noise surveys, particularly when the geometry of the survey is such that the distribution of the ambient noise sources favours the creation of the precursor waves over the direct wave Green's functions. If the array did not have the density of the one used in this study, the identification of the scattering and the location

of the scatterer may not have been possible. Without careful analysis, or with a sparse array, scattered arrivals could be mistaken for direct waves, and hence the inferred velocities would be incorrect. Passive scatterers join a list of effects (gaps in azimuthal sources, active sources, etc.) that cause arrivals in ambient noise surveys that can be misidentified as direct waves.

ACKNOWLEDGMENTS

We thank Fan-Chi Lin at Caltech for providing many useful suggestions and thank Dunzhu Li at Caltech for providing the finite-difference code. We also thank two anonymous reviewers for their helpful comments. Contribution number 212 from the Tectonics Observatory at Caltech. This work is supported by NSF (EAR-1045683) and the Gordon and Betty Moore Foundation.

REFERENCES

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.,* **169,** 1239–1260.

Brenguier, F., Shapiro, N.M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O. & Nercessian, A., 2008. Towards forecasting volcanic eruptions using seismic noise, *Nat. Geosci.,* **1,** 126–130.

Brenguier, F., Shapiro, N.M., Campillo, M., Nercessian, A. & Ferrazzini, V., 2007. 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations, *Geophys. Res. Lett.,* **34,** doi:10.1029/2006GL028586.

Hasselmann, K., 1963. A statistical analysis of the generation of microseisms, *Rev. Geophys.,* **1,** 177–210.

Kawakatsu, H., Ohminato, T. & Ito, H., 1994. 10s-period volcanic tremors observed over a wide area in Southwestern Japan, *Geophys. Res. Lett.,* **21,** 1963–1966.

Lin, F.C., Moschetti, M.P. & Ritzwoller, M.H., 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.,* **173,** 281–298.

Lin, F.C., Ritzwoller, M.H. & Shen, W.S., 2011. On the reliability of attenuation measurements from ambient noise cross-correlations, *Geophys. Res. Lett.,* **38,** doi:10.1029/2011GL047366.

Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.,* **110,** 3011–3017.

Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Philos. Trans. R. Soc. Lond., A,* **243,** 1–35.

Nagaoka, Y., Nishida, K., Aoki, Y., Takeo, M. & Ohminato, T., 2012. Seismic imaging of magma chamber beneath an active volcano, *Earth planet. Sci. Lett.,* **333,** 1–8.

Pritchard, M.E. & Simons, M., 2002. A satellite geodetic survey of large-scale deformation of volcanic centres in the central Andes, *Nature,* **418,** 167–171.

Pritchard, M.E. & Simons, M., 2004. An InSAR-based survey of volcanic deformation in the central Andes, *Geochem. Geophys. Geosyst.,* **5,** doi:10.1029/2003GC000610.

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31,** doi:10.1029/2004GL019491.

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science,* **307,** 1615–1618.

Shapiro, N.M., Ritzwoller, M.H. & Bensen, G.D., 2006. Source location of the 26 sec microseism from cross-correlations of ambient seismic noise, *Geophys. Res. Lett.,* **33,** L18310, doi:10.1029/2006GL027010.

Snieder, R., van Wijk, K., Haney, M. & Calvert, R., 2008. Cancellation of spurious arrivals in Green's function extraction and the generalized optical theorem, *Phys. Rev. E,* **78,** 036606, doi:10.1103/PhysRevE.78.036606.

Stehly, L., Campillo, M. & Shapiro, N.M., 2006. A study of the seismic noise from its long-range correlation properties, *J. geophys. Res. Solid Earth,* **111,** doi:10.1029/2005JB004237.

Tanimoto, T., 1990. Modeling curved surface-wave paths – membrane-surface wave synthetics, *Geophys. J. Int.,* **102,** 89–100.

Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.,* **93,** doi:10.1103/PhysRevLett.93.254301.

Wapenaar, K., Slob, E. & Snieder, R., 2010. On seismic interferometry, the generalized optical theorem, and the scattering matrix of a point scatterer, *Geophysics,* **75,** Sa27–Sa35.

Weeraratne, D.S., Forsyth, D.W., Fischer, K.M. & Nyblade, A.A., 2003. Evidence for an upper mantle plume beneath the Tanzanian craton from Rayleigh wave tomography, *J. geophys. Res.,* **108,** 2427, doi:10.1029/2002JB002273.

Yang, Y. & Ritzwoller, M.H., 2008. Characteristics of ambient seismic noise as a source for surface wave tomography, *Geochem. Geophys. Geosyst.,* **9,** doi:10.1029/2007GC001814.

Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.,* **168,** 259–274.

Yao, H., Campman, X., de Hoop, M.V. & van der Hilst, R.D., 2009. Estimation of surface wave Green's functions from correlation of direct waves, coda waves, and ambient noise in SE Tibet, *Phys. Earth planet. Inter.,* **177,** 1–11.

Yao, H., van der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I. Phase velocity maps, *Geophys. J. Int.,* **166,** 732–744.

Zeng, X.F. & Ni, S.D., 2010. A persistent localized microseismic source near the Kyushu Island, Japan, *Geophys. Res. Lett.,* **37,** doi:10.1029/2010GL045774.

Zeng, X. & Ni, S., 2011. Correction to "A persistent localized microseismic source near the Kyushu Island, Japan", *Geophys. Res. Lett.,* **38,** L16320, doi:10.1029/2011GL048822.

Zhan, Z.W., Ni, S.D., Helmberger, D.V. & Clayton, R.W., 2010. Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise, *Geophys. J. Int.,* **182,** 408–420.

Zheng, Y., Shen, W.S., Zhou, L.Q., Yang, Y.J., Xie, Z.J. & Ritzwoller, M.H., 2011. Crust and uppermost mantle beneath the North China Craton, northeastern China, and the Sea of Japan from ambient noise tomography, *J. geophys. Res. Solid Earth,* **116,** doi:10.1029/2011JB008637.